

# First-order convex feasibility algorithms for limited angular-range X-ray CT

Emil Y. Sidky\* and Xiaochuan Pan†

*University of Chicago*

*Department of Radiology*

*5841 S. Maryland Ave., Chicago IL, 60637*

Jakob H. Jørgensen‡

*Technical University of Denmark*

*Department of Informatics and Mathematical Modeling*

*Richard Petersens Plads, Building 321,*

*2800 Kgs. Lyngby, Denmark*

(Dated: September 6, 2012)

## Abstract

PURPOSE: Iterative image reconstruction (IIR) algorithms in Computed Tomography (CT) are based on algorithms for solving a particular optimization problem. Design of the IIR algorithm, therefore, is aided by knowledge of the solution to the optimization problem on which it is based. Often times, however, it is impractical to achieve accurate solution to the optimization of interest, which complicates design of IIR algorithms. This issue is particularly acute for CT with a limited angular-range scan, which leads to poorly conditioned system matrices and difficult to solve optimization problems. In this article, we develop IIR algorithms which solve a certain type of optimization called convex feasibility. The convex feasibility approach can provide alternatives to unconstrained optimization approaches and at the same time allow for efficient algorithms for their solution – thereby facilitating the IIR algorithm design process.

METHOD: An accelerated version of the Chambolle-Pock (CP) algorithm is adapted to various convex feasibility problems of potential interest to IIR in CT. One of the proposed problems is seen to be equivalent to least-squares minimization, and two other problems provide alternatives to penalized, least-squares minimization.

RESULTS: The accelerated CP algorithms are demonstrated on a simulation of circular fan-beam CT with a limited scanning arc of  $144^\circ$ . The CP algorithms are seen in the empirical results to converge their respective convex feasibility problems.

CONCLUSION: Formulation of convex feasibility problems can provide a useful alternative to unconstrained optimization when designing IIR algorithms for CT. The approach is amenable to recent methods for accelerating first-order algorithms which may be particularly useful for CT with limited angular-range scanning. The present article demonstrates the methodology, and future work will illustrate its utility in actual CT application.

---

\* sidky@uchicago.edu

† xpan@uchicago.edu

‡ jakj@imm.dtu.dk

## I. INTRODUCTION

Iterative image reconstruction (IIR) algorithms in computed tomography (CT) are designed based on some form of optimization. When designing IIR algorithms to account for various factors in the CT model, the actual designing occurs usually at the optimization problem and not the individual processing steps of the IIR algorithm. Once the optimization problem is established, algorithms are developed to solve it. Achieving convergent algorithms is important, because they yield access to the designed solution of the optimization problem and allow for direct assessment of what factors to include in a particular optimization. Convergent algorithms can also aid in determining at what iteration number to truncate an IIR algorithm. With access to the designed solution, the difference between it and previous iterates can be quantitatively evaluated to see whether this difference is significant with respect to a given CT imaging task.

It can be challenging to develop convergent algorithms for some optimizations problems of interest. This issue is particularly acute for CT, which involves large-scale optimization. In using the term “large-scale”, we are specifically referring to optimization problems based on a linear data model, and the dimension of the linear system is so large that the system matrix cannot be explicitly computed and stored in memory. Such systems only allow for algorithms which employ operations of a similar computational expense to matrix-vector products. Large-scale optimization algorithms are generally restricted to first-order methods, where only gradient information on the objective is used, or row-action algorithms such as the algebraic reconstruction technique (ART) [1, 2]. Recently, there has been renewed interest in developing convergent algorithms using  $\ell_1$ -based image norms, and only in the last couple of years have practical, convergent algorithms been developed to solve these optimization problems for IIR in CT [3–6]. Despite the progress in algorithms, there are still CT configurations of practical interest, which can lead to optimization problems that can be quite challenging to solve accurately. Of particular interest in this work is CT with a limited angular-range scanning arc. Such a configuration is relevant to many C-arm CT and tomosynthesis applications. Modeling limited angular-range scanning, leads to system matrices with unfavorable singular value spectra and optimization problems which are difficult to solve efficiently.

In this article, we consider application of convex feasibility [7, 8] to IIR for CT. In convex

feasibility, various constraints on properties of the image are formulated so that each of these constraints specifies a convex set. Taking the intersection of all of the convex sets yields a single convex set, and the idea is to simply choose one of these images in the intersection set. We have found convex feasibility to be useful for CT IIR algorithm design [9], and it is of particular interest here for limited angular-range CT, because convex feasibility is amenable to recent accelerated first-order algorithms proposed by Chambolle and Pock (CP) [10]. In Sec. II, we specify the limited angular-range CT system, discuss unconstrained optimization approaches, and then list three useful convex feasibility problems along with a corresponding accelerated CP algorithm. In Sec. III, the accelerated convex feasibility CP algorithms are demonstrated with simulated CT projection data.

## II. METHODS: CHAMBOLLE-POCK ALGORITHMS FOR CONVEX FEASIBILITY

For this article, we focus on modeling circular, fan-beam CT with a limited scanning angular range. As with most work on IIR, the data model is discrete-to-discrete (DD) and can be written as a linear equation:

$$\mathbf{g} = X\mathbf{f}, \quad (1)$$

where  $\mathbf{f}$  is the image vector comprised of pixel coefficients,  $X$  is the system matrix generated by computing the ray-integrals with the line-intersection method, and  $\mathbf{g}$  is the data vector containing the estimated projection samples. For the present investigation on IIR algorithms, we consider a single configuration for limited angular range scanning where the system matrix  $X$  has a left-inverse ( $X^T X$  is invertible) but is numerically unstable in the sense that it has a large condition number. The vector  $\mathbf{f}$  consists of the pixels within a circle inscribed in a  $256 \times 256$  pixel array; the total number of pixels is 51,468. The sinogram contains 128 views spanning a  $144^\circ$  scanning arc, and the projections are taken on a 512-bin linear detector array. The modeled source-to-isocenter and source-to-detector distances are 40 and 80 cm, respectively. The total number of transmission measurements is 65,536, and as a result the system matrix  $X$  has about 25% more rows than columns. The condition of  $X$ , however, is poor, which can be understood by considering the corresponding continuous-to-continuous (CC) fan-beam transform. A sufficient angular range for stable inversion of the CC fan-beam transform requires a  $208^\circ$  scanning arc ( $180^\circ$  plus the fan-angle). By using the inverse

power method, as described in Ref. [11], the condition number, the ratio of the largest to smallest singular value, for  $X$  is determined to be  $2.55 \times 10^4$ . One effect of the large condition number is to amplify noise present in the data, but it also can cause slow convergence for optimization-based IIR.

105 Image reconstruction using this DD data model is usually performed with some form of optimization, because physical factors and inaccuracy of the model render Eq. (1) inconsistent – namely, no  $\mathbf{f}$  exists satisfying this equation. Typically in using this model, quadratic optimizations are formulated, the simplest of which is the least-squares problem

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g} - X\mathbf{f}\|_2^2 \right\}, \quad (2)$$

110 where  $\mathbf{f}^*$  is the image which minimizes the Euclidean distance between the available data  $\mathbf{g}$  and the estimated data  $X\mathbf{f}$ . Taking the gradient of this objective, and setting it to zero component-wise, leads to the following consistent linear equation:

$$X^T X \mathbf{f} = X^T \mathbf{g}, \quad (3)$$

115 where the superscript  $T$  denotes the matrix transpose. This linear equation is particularly useful for setting up the linear conjugate gradients (CG) algorithm, see for example Ref. [12], which has been used as the gold standard algorithm for large-scale quadratic optimization in IIR. The reader is also referred to CGLS and LSQR, which solve Eq. (2) for non-symmetric  $X$  [13].

The solution to Eq. (2) or (3) can be undesirable because of inconsistency in the data. 120 Particularly for the present case, the poor conditioning of  $X$  can yield tremendously amplified artifacts in the reconstructed image. As is well-known, artifacts due to data inconsistency can be controlled in optimization-based IIR by adding a penalty term to discourage large variations between neighboring pixels:

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g} - X\mathbf{f}\|_2^2 + \alpha R(\mathbf{f}) \right\}, \quad (4)$$

125 where  $R(\mathbf{f})$  is a generic roughness term which usually is a convex function of the difference between neighboring pixels in the image. The parameter  $\alpha$  controls the strength of the penalty with larger values leading to smoother images. When  $R(\mathbf{f})$  is chosen to be quadratic in the pixel values, the optimization problem can be solved by a host of standard algorithms including CG. Of recent interest have been convex regularizers based on the  $\ell_1$ -norm, which

is more difficult to treat and, accordingly, for which many new, convergent algorithms have been proposed and applied to image reconstruction in CT [3–6].

In this article, we consider convex feasibility problems which provide alternatives to the above-mentioned optimization problems. For convex feasibility problems, convex sets resulting from constraints on various properties of the image are formulated, and a single image which satisfies all the imposed constraints is sought. Most algorithms for such problems are based on projection onto convex sets (POCS) [8], where the image estimate is sequentially projected on to each constraint set. Convex feasibility problems can be: inconsistent, no image satisfies all the constraints; or consistent, at least one image satisfies all the constraints. In either case, POCS algorithms can yield a useful solution. In the inconsistent case, POCS algorithms can be designed to yield an image “close” to satisfying all the constraints. In the consistent case, a POCS algorithm can be designed to find an image obeying all the constraints. In either case, the issue of uniqueness is secondary, as an image “in the middle” of many inconsistent constraints or in the intersection set of consistent constraints is considered to be equally valid. Accordingly, the POCS result often depends on starting image, relaxation schemes, and projection order.

For our purposes we write a general convex feasibility as the following optimization:

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \sum_i \delta_{S_i}(K_i(\mathbf{f})) \right\}, \quad (5)$$

$K_i(\cdot)$  is the  $i$ th affine transform of the image  $\mathbf{f}$ ;  $S_i$  is the  $i$ th convex set to which  $K_i(\mathbf{f})$  belongs; and the indicator function  $\delta$  is defined

$$\delta_S(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in S \\ \infty & \mathbf{x} \notin S \end{cases}. \quad (6)$$

The use of indicator functions in convex analysis provides a means to turn convex sets into convex functions [14], and in this case, they allow convex feasibility problems to be written as a minimization of a single objective function. The objective function in Eq. (5) is zero for any image  $\mathbf{f}$  satisfying all the constraints, i.e.  $K_i(\mathbf{f}) \in S_i$  for all  $i$ , and it is infinity if any of the constraints are violated. For a consistent convex feasibility problem, the objective minimum is zero, and for an inconsistent convex feasibility problem, the objective minimum is infinity.

To solve the generic convex feasibility problem in Eq. (5), we modify this optimization by adding a quadratic term:

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \sum_i \delta_{S_i}(K_i(\mathbf{f})) \right\}, \quad (7)$$

where  $\mathbf{f}_{\text{prior}}$  is a prior image estimate that can be set to zero if no prior image is available. With this optimization problem, we actually specify a unique solution to our generic convex feasibility problem in the consistent case – namely the image satisfying all constraints and closest to  $\mathbf{f}_{\text{prior}}$ . As we will demonstrate the algorithm we propose to use for solving Eq. (7) appears to yield useful solutions for the *inconsistent* case. This latter property can be important for IIR in CT because the data model in Eq. (1) is often inconsistent with the available projection data.

The reason for recasting the optimization in the form shown in Eq. (7) is that this optimization can be solved by an accelerated algorithm described in Ref. [10]. Recently, we have been interested in a convex optimization framework and algorithms derived by Chambolle and Pock (CP) [10, 15]. In a previous publication [6], we illustrated how to use Algorithm 1 from Ref. [10] to prototype many optimization problems of potential interest to image reconstruction in CT. In the present work, we narrow the class of optimization problems to those which can be written in the form of Eq. (7), where the sets  $S_i$  are simple enough that a direct Euclidean projection to the  $S_i$  is analytically available. Considering this form of optimization has the algorithmic advantage that the objective can be split between a smooth, strongly convex term, the first quadratic, and non-smooth indicator functions. Because of the smooth term, Algorithm 2 from Ref. [10] can be employed which exploits acceleration for first-order algorithms developed by Nesterov [16, 17]. Accordingly we refer to Algorithm 2 in Ref. [10] as the accelerated CP algorithm. Algorithm acceleration is particularly important for IIR involving an ill-conditioned data model such as Eq. (1) in the case of limited angular range scanning.

In the following, we write various imaging problems in the form of Eq. (7). We consider the following three convex feasibility problems: CF1, one set specifying a data equality constraint; CF2, one set specifying a data inequality constraint; and CF3, two sets specifying data and total variation (TV) inequality constraints. The derived, accelerated CP instance for each problem is labeled CF1-CP, CF2-CP, and CF3-CP, respectively. Using simulated fan-beam CT data with limited angular-range scanning arc, Sec. III presents results for

all three problems in the consistent case and problems CF1 and CF3 in the inconsistent  
 190 case. Of particular importance, CF1-CP applied to the inconsistent case appears to solve  
 the ubiquitous least-squares optimization efficiently in the sense that its convergence is  
 competitive with CG.

### A. CF1-CP: an accelerated CP instance for a data equality constraint

The data model in Eq. (1) cannot be used directly as an implicit imaging model for  
 195 real CT data, because inconsistencies inherent in the data prevent a solution. But treating  
 this equation as an implicit imaging model for ideal simulation can be useful for algorithm  
 comparison and testing implementations of the system matrix  $X$ ; we use it for the former  
 purpose. We write this ideal imaging problem into an instance of Eq. (7):

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \delta_0(X\mathbf{f} - \mathbf{g}) \right\}, \quad (8)$$

200 where the indicator  $\delta_0(\cdot)$  is zero only when all components of the argument vector are zero,  
 and otherwise it is infinity. In matching this optimization with Eq. (7), there is only one  
 convex constraint where  $K_1(\mathbf{f}) = X\mathbf{f} - \mathbf{g}$  and  $S_1$  is the 0-vector with size,  $\text{size}(\mathbf{g})$ . In  
 considering ideal data and a left-invertible system matrix  $X$ , there is only one image for  
 which the indicator is not infinite. In this situation, the first quadratic has no effect on the  
 205 solution and accordingly the solution is independent of the prior image estimate  $\mathbf{f}_{\text{prior}}$ . If the  
 system matrix is not left-invertible, the solution to Eq. (8) is the image satisfying Eq. (1)  
 closest to  $\mathbf{f}_{\text{prior}}$ .

Following the formalism of Ref. [10], we write an accelerated CP instance for solving  
 Eq. (8) in Fig. 1. The algorithm is primal-dual, meaning that it solves simultaneously the  
 210 primal minimization, in this case Eq. (8), and a dual maximization. The combination of  
 solving the two optimizations allows for a convergence check by computing the primal-dual  
 gap. In this article, for brevity, we do not write the dual maximization nor explain this check  
 further, but we have covered it extensively in Ref. [6]. We do, however, point out when this  
 check has been used in the present text. We define the pseudocode variables and operations  
 215 starting from the first line. The variable  $L$  is assigned the matrix  $\ell_2$ -norm of  $X$ , which is the  
 largest singular value. This quantity can be computed by the standard power method, see  
 [6] for its application in the present context. The parameters  $\tau$  and  $\sigma$  control the step sizes



---

```

1:  $L \leftarrow \|X\|_2$ ;  $\tau \leftarrow 1$ ;  $\sigma \leftarrow 1/L^2$ ;  $n \leftarrow 0$ 
2: initialize  $\mathbf{f}_0$  and  $\mathbf{y}_0$  to zero vectors
3:  $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$ 
4: repeat
5:    $\mathbf{y}_{n+1} \leftarrow \mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})$ 
6:    $\mathbf{f}_{n+1} \leftarrow [\mathbf{f}_n - \tau(X^T\mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}})] / (1 + \tau)$ 
7:    $\theta \leftarrow 1/\sqrt{1 + 2\tau}$ ;  $\tau \leftarrow \tau\theta$ ;  $\sigma \leftarrow \sigma/\theta$ 
8:    $\bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$ 
9:    $n \leftarrow n + 1$ 
10: until  $n \geq N$ 

```

---

**FIG. 1:** Pseudocode for  $N$ -steps of the accelerated CP algorithm instance for solving Eq. (8). Variables are defined in the text.

in the primal and dual problems respectively, and they are initialized so that their product yields  $1/L^2$ . Other choices on how to balance the starting values of  $\tau$  and  $\sigma$  can be made, but we have found that the convergence of our examples does not depend strongly on the initial choice of these parameters. Line 5 shows the update of the dual variable  $\mathbf{y}_{n+1}$ ; this variable has the same dimension as the data vector  $\mathbf{g}$ . Line 6 updates the image, and Line 7 adjusts the step-sizes in a way that accelerates the CP algorithm [10].

## B. CF2-CP: an accelerated CP instance for inequality constrained data-error

Performing IIR with projection data containing inconsistency, requires some form of image regularization. One common strategy is to employ Tikhonov regularization, see for example Chapter 2 of Ref. [18]. Tikhonov regularization fits into the form of Eq. (4) by writing  $R(\mathbf{f}) = (1/2)\|\mathbf{f}\|_2^2$ . One small inconvenience with this approach, however, is that the physical units of the two terms in the objective of Eq. (4) are different, and therefore it can be difficult to physically interpret the regularization parameter  $\alpha$ . An equivalent

---

```

1:  $L \leftarrow \|X\|_2$ ;  $\tau \leftarrow 1$ ;  $\sigma \leftarrow 1/L^2$ ;  $n \leftarrow 0$ 
2: initialize  $\mathbf{f}_0$  and  $\mathbf{y}_0$  to zero vectors
3:  $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$ 
4: repeat
5:    $\mathbf{y}_{n+1} \leftarrow \max(\|\mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})\|_2 - \sigma\epsilon, 0) (\mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g}))$ 
6:    $\mathbf{f}_{n+1} \leftarrow [\mathbf{f}_n - \tau(X^T \mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}})] / (1 + \tau)$ 
7:    $\theta \leftarrow 1/\sqrt{1 + 2\tau}$ ;  $\tau \leftarrow \tau\theta$ ;  $\sigma \leftarrow \sigma/\theta$ 
8:    $\bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$ 
9:    $n \leftarrow n + 1$ 
10: until  $n \geq N$ 

```

---

**FIG. 2:** Pseudocode for  $N$ -steps of the accelerated CP algorithm instance for solving Eq. (9) with parameter  $\epsilon$ . Variables are defined in Sec. II A.

optimization can be formulated as a special case of Eq. (7):

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \delta_{\text{Ball}(\epsilon)}(X\mathbf{f} - \mathbf{g}) \right\}, \quad (9)$$

which differs from Eq. (8) only in that the set  $S_1$  is widened from a 0-vector to  $\text{Ball}(\epsilon)$ , where we use the term  $\text{Ball}(\epsilon)$  to denote a multi-dimensional solid sphere of radius  $\epsilon$  (the dimension of the solid sphere is taken to be the same as  $\text{size}(\mathbf{g})$ ). The indicator  $\delta_{\text{Ball}(\epsilon)}(X\mathbf{f} - \mathbf{g})$  is zero when  $\|X\mathbf{f} - \mathbf{g}\|_2 \leq \epsilon$  and infinity otherwise. This optimization is equivalent to Tikhonov regularization when  $\mathbf{f}_{\text{prior}}$  is zero and  $\epsilon > 0$  in the sense that there exists a corresponding  $\alpha$  (not known ahead of time) where the two optimizations yield the same solution. The advantage of Eq. (9) is that the parameter  $\epsilon$  has a meaningful physical interpretation as a tolerance on the data-error. Larger  $\epsilon$  yields greater regularization. Generally, the Tikhonov form is preferred due to algorithm availability. Tikhonov regularization can be solved efficiently, for example, by linear CG. With the application of CF2-CP, however, an accelerated solver is now available that directly solves the constrained, minimization in Eq. (9).

The pseudocode for CF2-CP is given in Fig. 2. This pseudocode differs from the previous at the update of the dual variable  $\mathbf{y}_{n+1}$  in Line 5. The derivation of this dual update is

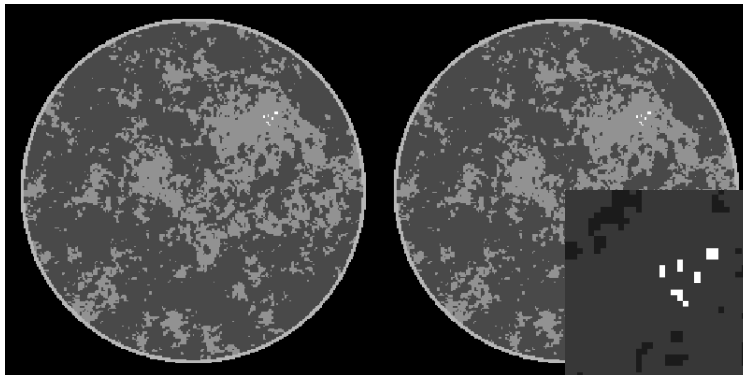
covered in detail in our previous work on the application of the CP algorithm to CT image reconstruction [6]. For the limited angular-range CT problem considered here, Eq (9) is particularly challenging because the constraint shape is highly eccentric due to the spread  
 250 in singular values of  $X$ .

### C. CF3-CP: convex feasibility for total variation and data-error constraints

Recently, regularization based on the  $\ell_1$ -norm has received much attention. In particular, the TV semi-norm has found extensive application in medical imaging due to the fact that CT tomographic images are approximately piece-wise constant. The TV semi-norm of  $\mathbf{f}$  is  
 255 written as  $\|(|\nabla \mathbf{f}|)\|_1$ , where  $\nabla$  is a matrix encoding a finite-difference approximation to the gradient operator; it acts on an image and yields a spatial-vector image. The absolute value operation acts pixel-wise, taking the length of the spatial-vector at each pixel of this image; accordingly,  $|\nabla \mathbf{f}|$  is the gradient-magnitude image of  $\mathbf{f}$ . The TV semi-norm can be used as a penalty with the generic optimization of Eq. (4), by setting  $R(\mathbf{f}) = \|(|\nabla \mathbf{f}|)\|_1$ . Efficient  
 260 large scale solvers for this optimization problem have only recently been developed with some algorithms relying on smoothing the TV term [3–5]. As with Tikhonov regularization, there is still the inconvenience of having no physical meaning of the regularization parameter  $\alpha$ . We continue along the path of recasting optimization problems as a convex feasibility problem and consider the following:

$$265 \quad \mathbf{f}^* = \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \delta_{\text{Ball}(\epsilon)}(X\mathbf{f} - \mathbf{g}) + \delta_{\text{Diamond}(\gamma)}(|\nabla \mathbf{f}|) \right\}, \quad (10)$$

where the additional indicator places a constraint on the TV of  $\mathbf{f}$ ; and we have  $K_1(\mathbf{f}) = X\mathbf{f} + \mathbf{g}$ ,  $K_2(\mathbf{f}) = \nabla \mathbf{f}$ ,  $S_1 = \{\mathbf{g} \text{ such that } \mathbf{g} \in \text{Ball}(\epsilon)\}$ , and  $S_2 = \{\mathbf{z} \text{ such that } |\mathbf{z}| \in \text{Diamond}(\gamma)\}$ , where  $\mathbf{z}$  is a spatial-vector image. The term  $\text{Diamond}(\gamma)$  describes the  $\ell_1$ -ball of scale  $\gamma$ ; the indicator  $\delta_{\text{Diamond}(\gamma)}(|\nabla \mathbf{f}|)$  is zero when  $\|(|\nabla \mathbf{f}|)\|_1 \leq \gamma$ . This convex feasibility problem asks  
 270 for the image that is closest to  $\mathbf{f}_{\text{prior}}$  and satisfies the  $\epsilon$ -data-error and  $\gamma$ -TV-constraints. We demonstrate in Sec. III application of CF3-CP to both inconsistent and consistent constraint sets. Due to the length of the pseudocode, we present it in Appendix A, and point out that it can be derived following Ref. [6], using the Moreau identity described in Ref. [10] and an algorithm for projection onto the  $\ell_1$ -ball [19].



**FIG. 3:** Breast phantom for the CT limited angular-range scanning simulation. Left: the phantom in the gray scale window  $[0.95, 1.15]$ . Right: the same phantom with a blow-up on the micro-calcification ROI displayed in the gray scale window  $[0.9, 1.8]$ . The right panel is the reference for all image reconstruction algorithm results.

We demonstrate the application of the various accelerated CP algorithm instances on simulated CT data generated from the breast phantom shown in Fig. 3. The phantom, described in Ref. [20, 21], is digitized on a  $256 \times 256$  pixel array. Four tissue types are modeled: the background fat tissue is assigned a value of 1.0, the modeled fibro-glandular tissue takes a value of 1.1, the outer skin layer is set to 1.15, and the micro-calcifications are assigned values in the range  $[1.8, 2.3]$ . The simulated CT configuration is described at the beginning of Sec. II.

In the following, the IIR algorithms are demonstrated with ideal data generated by applying the system matrix  $X$  to the phantom and with inconsistent data obtained by adding Poisson distributed noise to the ideal data set. We emphasize that the goal of the paper is to address convergence of difficult optimization problems related to IIR in limited angular-range CT. Thus, we are more interested in establishing that the CP algorithm instances achieve accurate solution to their corresponding optimization problems, and we are less concerned about the image quality of the reconstructed images. In checking convergence in the consistent case, we can monitor the primal-dual gap. For the inconsistent case, we hypothesize that CF1-CP minimizes the least-squares problem, Eq. (2), and we can use the gradient magnitude of the least-squares objective to check this hypothesis and test convergence; CF2-CP, we hypothesize, solves the same problem in the inconsistent case, and is

not interesting because we can instead use the parameter-less CF1 problem; and finally, for CF3-CP we do not have a convergence check in the inconsistent case and instead we monitor other metrics of the IIR. Two such metrics, which we use, are image root mean square error (RMSE) and data RMSE. The image RMSE is:

$$\frac{\|\mathbf{f} - \mathbf{f}_{\text{phantom}}\|_2}{\sqrt{\text{size}(\mathbf{f})}},$$

and the data RMSE is:

$$\frac{\|\mathbf{g} - X\mathbf{f}\|_2}{\sqrt{\text{size}(\mathbf{g})}}.$$

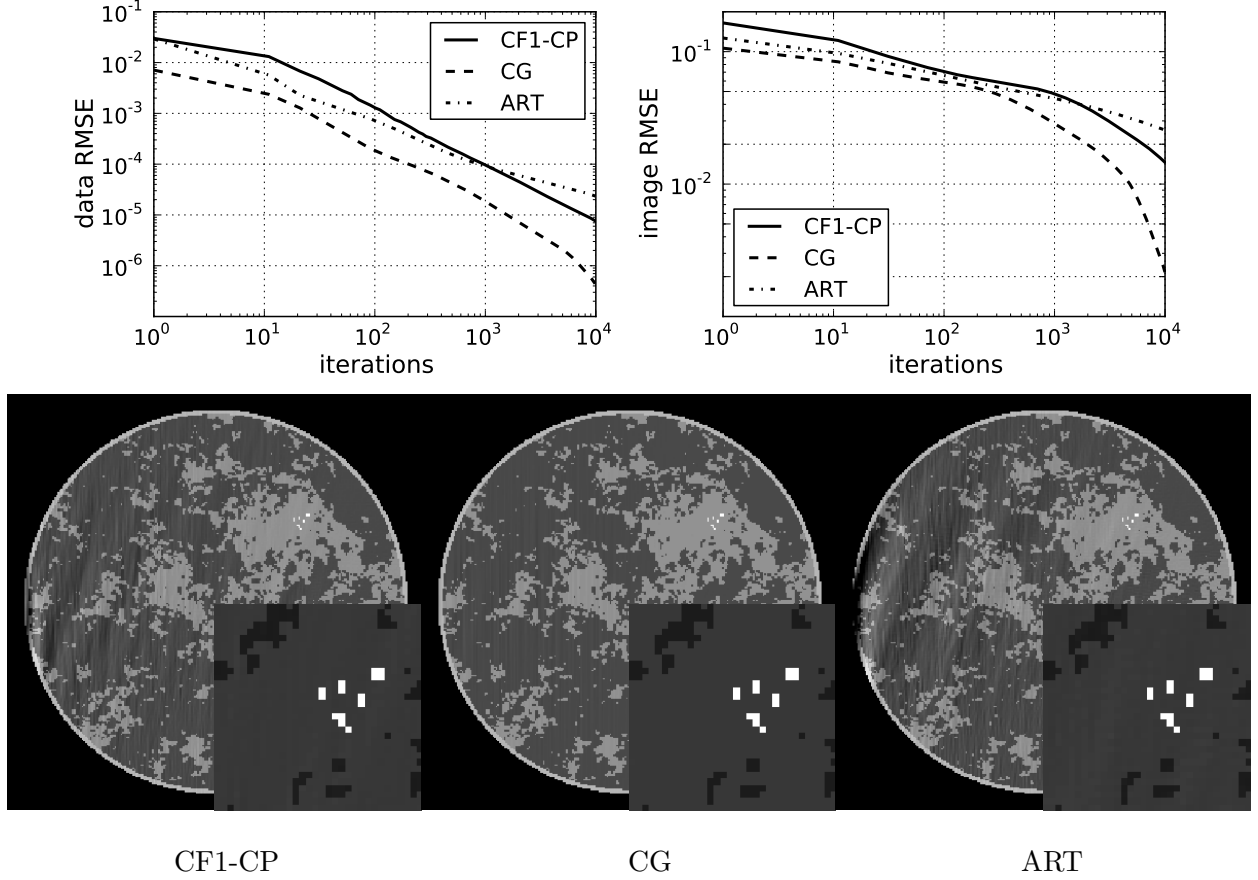
### A. Ideal data and equality-constrained optimization

We generate ideal data from the breast phantom and apply CF1-CP, with  $\mathbf{f}_{\text{prior}} = 0$ , to investigate its convergence behavior for limited angular-range CT. As the simulations is set up so that  $X$  is left-invertible and the data are generated from applying this system matrix to the test phantom, the indicator  $\delta_0(X\mathbf{f} - \mathbf{g})$  in Eq. (8) is zero only when  $\mathbf{f}$  is the phantom. It is the convergence to the breast phantom which is of interest here.

In order to have a reference to standard algorithms, we apply linear CG [12] and the ART to the same problem. Linear CG solves the minimization in Eq. (2), which corresponds to solving the linear system in Eq. (3). The matrix,  $X^T X$ , in this equation is symmetric with non-negative singular values. The ART algorithm, which is a form of POCS, solves Eq. (1) directly by cycling through orthogonal projections onto the hyper-planes specified by each row of the linear system.

The results of each algorithm are shown in Fig. 4. As the data are ideal, each algorithm drives the data-error to zero. The linear CG algorithm shows the smallest data RMSE, but we note similar slopes on the log-log plot of CG and CF1-CP during most of the computed iterations except near the end, where the slope of the CG curve steepens. The ART algorithm reveals a convergence slightly faster than CF1-CP, initially, but it is overtaken by CF1-CP near iteration 1000.

Because  $X$  is designed to be left-invertible, we know also that the image estimates must converge to the breast phantom for each of the three algorithms. A similar ordering of the convergence rates is observed in the image RMSE plot, but we note that the values of the image RMSE are all much larger than corresponding values in the data RMSE plots. This



**FIG. 4:** Results of CF1-CP with ideal, simulated data. Convergence is also compared with linear CG and ART. Top row: (Left) convergence of the three algorithms in terms of data RMSE, and (Right) convergence of the three algorithms in terms of image RMSE. Bottom row: the image at iteration 10000 for CF1-CP, CG, and ART shown in the same gray scale as Fig. 3. The artifacts seen at the right of the image and the large image RMSE are indications of the poor conditioning of  $X$ .

stems from the poor conditioning of  $X$ , and this point is emphasized in examining the shown image estimates at iteration 10000 for each algorithm.

While the image RMSE gives a summary metric on the accuracy of the image reconstruction, the displayed images at the final computed iteration number yields more detailed information on the image error incurred by truncating the algorithm iteration. The CF1-CP and ART images show wavy artifacts on the left side; the limited-angle scanning arc is over the right-side of the object. But the CG image shows visually accurate image reconstruction at the given gray scale window setting.

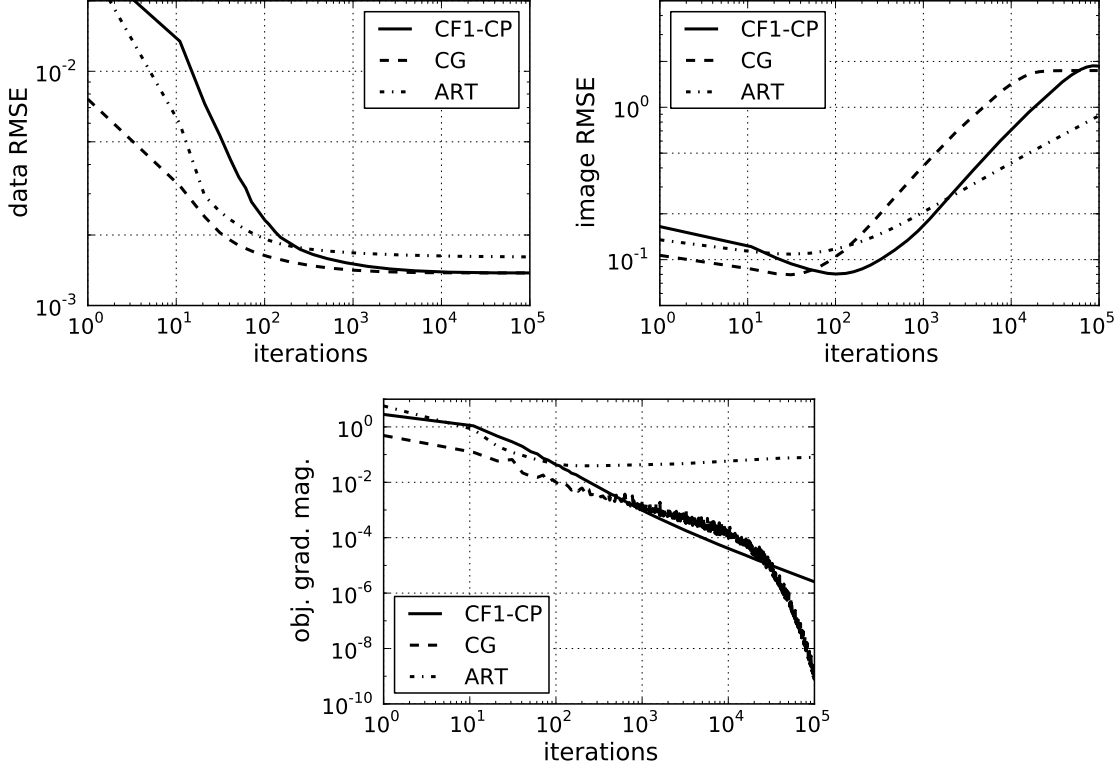
This initial result shows promising convergence rates for CF1-CP and that it may be competitive with existing algorithms for solving large, consistent linear systems. But we cannot draw any general conclusions on algorithm efficiency, because different simulation conditions may yield different ordering of the convergence rates. Moreover, we have implemented only the basic forms of CG and ART; no attempt at pre-conditioning CG was made and the relaxation parameter of ART was fixed at 1.

In this article, because we are interested in algorithm convergence, we discuss convergence criteria for each CP instance. For the present consistent problem, the primal-dual gap does provide a useful check of convergence for CF1-CP and for the present simulation this was monitored. Due to the special circumstances of this ideal simulation, data and image RMSE provide sufficient checks on convergence. For each algorithm, data and image RMSE are seen to decrease monotonically.

## B. Noisy, inconsistent data and equality-constrained optimization

In this section, we repeat the previous simulation with all three algorithms except that the data now contain inconsistency modeling Poisson distributed noise. The level of the noise is selected to simulate what could be seen in a low-dose CT scan. The use of this data model contradicts the application of equality-constrained optimization and CF1 becomes inconsistent. Before applying CF1-CP, we know that the indicator, and hence the whole objective, in Eq. (8) will have the value of infinity. But nothing prevents us from executing the CF1-CP operations, and accordingly we do so in this section. The linear CG algorithm can still be applied in this case, because the optimization in Eq. (2) is still well-defined even though there is no  $\mathbf{f}$  such that  $\mathbf{g} = X\mathbf{f}$ . Likewise, the linear system in Eq. (3) does have a solution even when  $\mathbf{g}$  is inconsistent. The basic ART algorithm, like CF1-CP, is not suited to this data model, because it is a solver for Eq. (1), which we know ahead of time has no solution. Again, like CF1-CP, the steps of ART can still be executed even with inconsistent data, and we show the results here.

In Fig. 5, we show evolution plots of quantities derived from the image estimates from each of the three algorithms. Because the data are inconsistent, the data- and image-error plots have a different behavior than the previous consistent example. In this case, we know that the data RMSE cannot be driven to zero, and each of the three algorithms seem to be



**FIG. 5:** Results of CF1-CP with noisy and inconsistent, simulated data. Convergence is also compared with linear CG and ART. Top row: (Left) evolution of the three algorithms in terms of data RMSE, and (Right) evolution of the three algorithms in terms of image RMSE. Bottom row: Convergence of the gradient of the quadratic objective in Eq. (2).

converging on a value greater than zero. The image RMSE shows an initial decrease to some minimum value followed by an upward trend. For CG the upward trend begins to level off at 20,000 iterations, while for CF1-CP this happens shortly before the final 100,000th iteration. The last plot in this figure shows the magnitude of the gradient of the least-squares objective from Eq. (2). The curve for the CG results shows an overall convergence by this metric, as expected, because this algorithm is designed to solve the normal equations of Eq. (2). The ART algorithm shows an initial decay followed by a slow increase. Again this result is expected, because ART is designed to solve Eq. (1) directly and not the least-squares minimization in Eq. (2). As an aside, we point at that in applying ART to inconsistent data it is important to allow the relaxation parameter to decay to zero. Interestingly, the CF1-CP algorithm shows a monotonic decrease of this gradient.

The result obtained by CF1-CP is quite surprising because the CP algorithm is primal-



375 dual. Without getting into too much mathematical detail, a consequence of inconsistent data  
being used for CF1-CP is that the dual maximization to Eq. (8) cannot converge. Indeed,  
the magnitude of the dual variable  $\mathbf{y}_n$  in Fig. 1 increases steadily with iteration number.  
Even though the dual problem diverges, this simulation seems to indicate convergence of  
the primal least-squares minimization in that the gradient of this objective is observed to  
380 monotonically decrease. And CF1-CP appears to achieve slightly lower data RMSE than  
linear CG. There is no proof that we are aware of, which covers this situation, thus we  
cannot claim that CF1-CP will always converge the least-squares problem. Therefore, in  
applying CF1-CP in this way it is crucial to evaluate the convergence criterion and to verify  
that the magnitude of the objective’s gradient decays to zero. The primal-dual gap cannot  
385 be used as a check for CF1-CP applied to inconsistent data.

The dependence of the gradient of the least-squares objective for the CF1-CP and CG  
algorithms is quite interesting. Between 10 and 20,000 iterations, CF1-CP shows a steeper  
decline in this convergence metric. But greater than 20,000 iterations the CG algorithm  
takes over and this metric drops precipitously. The CG behavior can be understood in  
390 realizing that the image has approximately 50,000 unknown pixel values and if there is  
no numerical error in the calculations, the CG algorithm terminates when the number of  
iterations equals the number of unknowns. Because numerical error is present, we do not  
observe exact convergence when the iteration number reaches 50,000, but instead the steep  
decline in the gradient of the least-squares objective is observed. This comparison between  
395 CF1-CP and CG has potential implications for larger systems where the steep drop-off for  
CG would occur at higher iteration number.

The conditions of this particular simulation are not relevant to practical application  
because it is already well-known that minimizing data-fidelity objectives with noisy data  
converges to an extremely noisy image particularly for an ill-conditioned system matrix;  
400 noting the large values of the image RMSE we know this to be the case without displaying  
the image. But this example is interesting in investigating convergence properties. While it  
is true that monitoring the gradient magnitude of the least-squares objective yields a sense  
about convergence, we do not know *a priori* what threshold this metric needs to cross before  
we can say the IIR is converged, see Ref. [22] for further discussion on this point related to  
405 IIR in CT. This example in particular highlights the point that an image metric of interest,  
such as image and data RMSE, needs to be observed to level off in combination with a

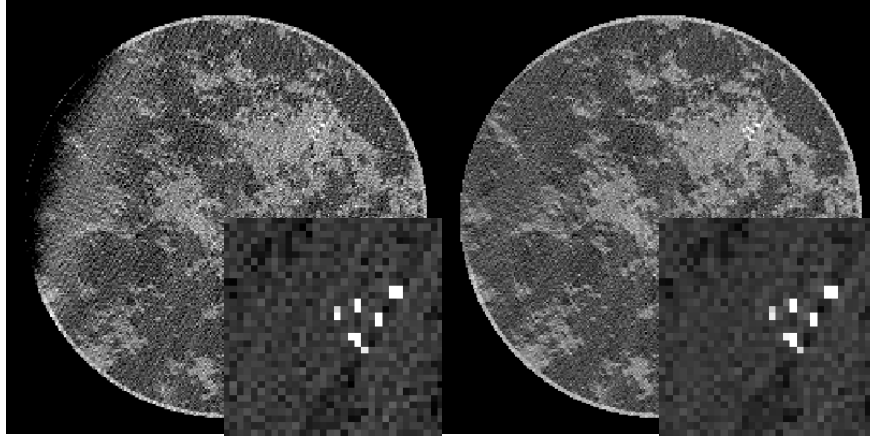
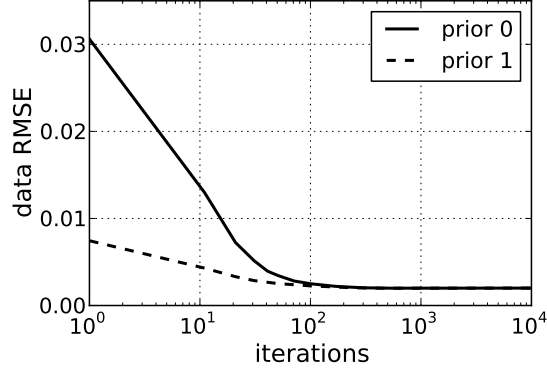
steady decrease of a convergence metric. For this example convergence of the image RMSE when the gradient of the least-squares objective drops below  $10^{-5}$ , while the data RMSE convergence occurs earlier.

### C. Noisy, inconsistent data with inequality-constrained optimization

In performing IIR with inconsistent projection data, some form of regularization is generally needed. In using the convex feasibility approach, we apply CF2-CP after deciding on the parameter  $\epsilon$ . The parameter  $\epsilon$  has a minimum value, below which no images satisfy the data-error constraint, and larger  $\epsilon$  leads to greater image regularity. The choice of  $\epsilon$  may be guided by properties of the available data or a prior reconstruction. In this case, we have results from the previous section and we note that the data RMSE achieve values below 0.002. Accordingly, for the present simulation we select a tight data-error constraint  $\epsilon = 0.512$ , which is equivalent to allowing a data RMSE of 0.002. The CF2-CP algorithm selects the image obeying the data-error constraint closest to  $\mathbf{f}_{\text{prior}}$ , and to illustrate the dependence on  $\mathbf{f}_{\text{prior}}$  we present results for two choices: an image of zero values, and an image set to 1 over the support of the phantom. Note that the second choice assumes prior knowledge of the object support and background value of 1. To our knowledge, there is no direct, existing algorithm for solving Eq. (9), and thus we display results for CF2-CP only. One can use a standard algorithm such as linear CG to solve the Lagrangian form of Eq. (9), but this method is indirect because it is not known ahead of time what Lagrange multiplier leads to the desired value of  $\epsilon$ .

The results of CF2-CP are shown in Fig. 6. The data RMSE is seen to converge to the value established by the choice of  $\epsilon$ . In the displayed images, there is a clear difference due to the choice of prior image. The image resulting from the zero prior shows a substantial drift of the gray level on the left side of the image. Application of a prior image consisting of constant background values over the object's true support removes this artifact almost completely. These results indicate that use of prior knowledge, when available, can have a large impact on image quality particularly for an ill-conditioned system matrix such as what arises in limited angular-range CT.

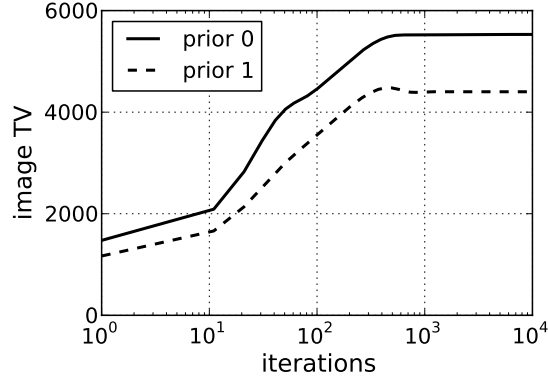
Convergence of the CF2-CP algorithm can be checked by the primal-dual gap, when  $\epsilon$  is chosen sufficiently large that at least one image  $\mathbf{f}$  satisfies the data-error constraint. We



**FIG. 6:** Results of CF2-CP with noisy and inconsistent, simulated data. The curves labeled “prior 0” correspond to a zero prior image. The curves labeled “prior 1” correspond to a prior image of 1.0 on the object support. Top: convergence of the data RMSE to the preset value of 0.002. Bottom: (Left) “prior 0” final image, and (Right) “prior 1” final image. Gray scales are the same as Fig. 3.

have performed this check for this simulation, but we do not display the primal-dual gap plot. While only  $\epsilon$  admitting at least one feasible image are of practical interest, we have observed that the CF2-CP algorithm appears to converge even for choices of  $\epsilon$  too small to admit a consistent  $\mathbf{f}$ , much like the use of CF1-CP on inconsistent data. We do not pursue the use of CF2-CP with inconsistent  $\epsilon$  further, because it does not have any use beyond what is available through the use of CF1-CP. As we have seen in the previous section, CF1-CP appears to be capable of finding the minimum consistent  $\epsilon$  as it is seen to minimize Eq. (2). An attractive feature of CF2-CP is that the rapid convergence seen for CF1-CP appears to carry over to CF2-CP. The image estimate reaches the desired data RMSE within 1000 iterations.

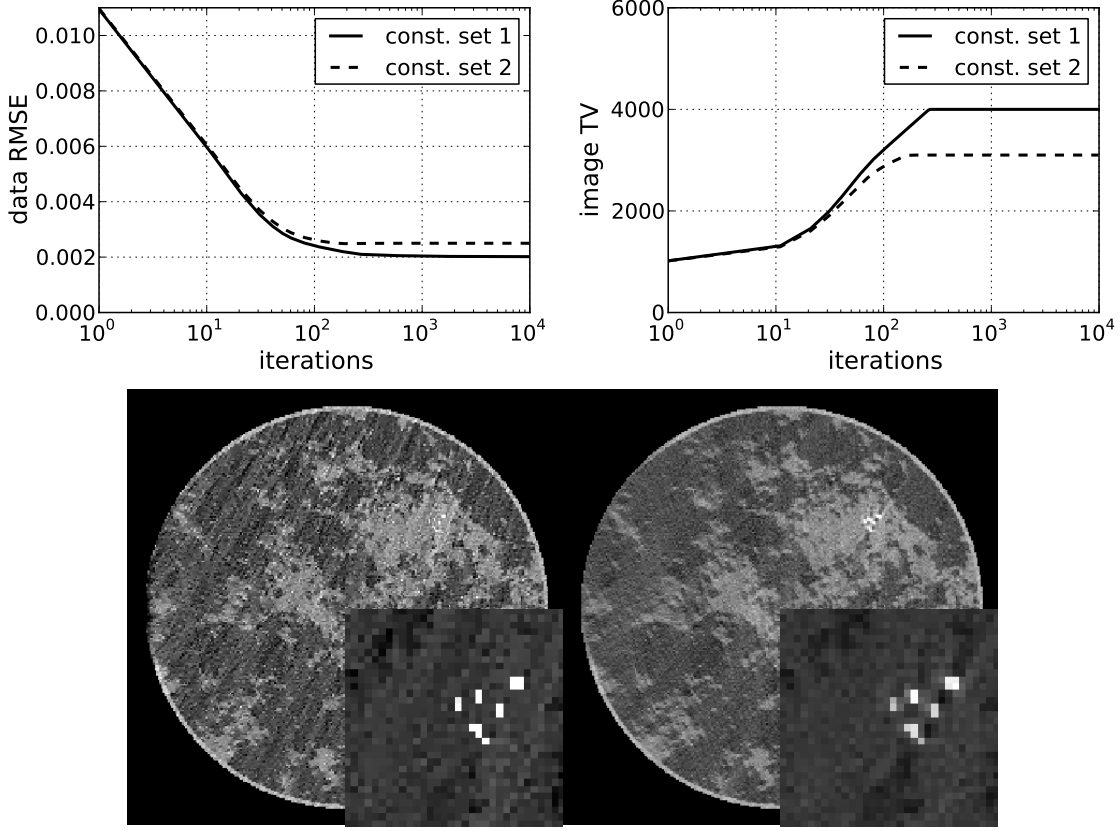
#### D. Noisy, inconsistent data with two-set convex feasibility



**FIG. 7:** Values of image TV for the results shown in Fig. 6. These values help to design the constraint values for Eq. (10).

For the last demonstration of the convex feasibility approach to IIR for limited-angular range CT, we apply CF3-CP, which seeks the image closest to a prior image and respects constraints on image TV and data-error. We are unaware of other algorithms, which address this problem, and only results for CF3-CP are shown. In applying CF3-CP, we need two constants,  $\epsilon$  and  $\gamma$ , and accordingly use of this algorithm is meant to be preceded by an initial image reconstruction in order to have a sense of interesting values for the data-error and image TV constraints. From the previous results, we already have information about data-error, and because we have the image estimates, we can also compute image TV values shown in Fig. (7). The image TV values corresponding to the two prior image estimates differ significantly, reflecting the quite different appearance of the resulting images shown in Fig. 6. We follow the use of the support prior image, and note that the computation settles on a value near 4400 for the image TV.

In our first example with this two-set convex feasibility problem, we maintain the tight data-error constraint  $\epsilon = 0.512$  (a data RMSE of 0.002) but attempt to find an image with lower TV by selecting  $\gamma = 4000$ . The results for these settings, labeled “const. set 1”, are shown in Fig. 8. Interestingly, this set of constraints appears to be just barely infeasible; the CF3-CP algorithm converges to an image TV of 4000.012 and a data RMSE of 0.00202. Furthermore, the dual variable magnitude increases steadily, an indication of an infeasible problem. The curves for image TV and data RMSE to appear to indicate some type of



**FIG. 8:** Results of CF3-CP with noisy and inconsistent, simulated data for two different constraint values: “const. set 1” refers to choosing  $\epsilon = 0.512$  (a data RMSE of 0.002) and  $\gamma = 4000$ ; “const. set 2” refers to choosing  $\epsilon = 0.768$  (a data RMSE of 0.0025) and  $\gamma = 3100$ . Top row: (Left) evolution of data RMSE, and (Right) evolution of image TV. Bottom row: (Left) resulting image of “const. set 1”, and (Right) resulting image of “const. set 2”. Gray scales are the same as Fig. 3.

convergence just as CF1-CP appears to converge to the solution of Eq. (2) for inconsistent data. But investigation of the convergence properties of CF3-CP for infeasible constraints is beyond the scope of this article.

In the second example, we loosen the data-error constraint to  $\epsilon = 0.768$  (a data RMSE of 0.0025) and seek an image with lower TV,  $\gamma = 3100$ , and the results are also shown in Fig. 8. In this case, the constraint values are met by CF3-CP, and the resulting image has noticeably less noise than the images with no TV constraint imposed shown in Fig. 6 particularly in the ROI containing the model micro-calcifications. Because this constraint set contains feasible solutions, the primal-dual gap can be used as a convergence check for

CF3-CP.

We note that for both choices of constraint sets, the convergence rate appears to be similar  
480 to CF1-CP and CF2-CP. Because CF1-CP compares favorably with CG, we speculate that  
CF3-CP solves the two-set convex feasibility problem with high efficiency. This efficiency  
is particularly important for IIR in limited angular-range CT, where the system matrix is  
poorly conditioned.

Although related to constrained, TV-minimization [6, 23] or TV-penalized, least-squares  
485 [3–6], CF3 is a different problem. Constrained, TV minimization and TV-penalized, least  
squares depend on a single parameter exploring some kind of optimal balance between image  
TV and data-error. Use of CF3, on the other hand, depends on two parameters allowing  
investigation of image quality for fixed data-error and varying image TV, or *vice versa*. In  
summary, CF3 provides an alternative design for TV regularized IIR.

#### 490 IV. CONCLUSION

We have illustrated three examples of convex feasibility problems for IIR applied to  
limited angular-range CT, which provide alternative designs to unconstrained optimization  
problems formulated for IIR in CT.

One of the motivations of the alternative design is that these convex feasibility problems  
495 are amenable to the accelerated CP algorithm, and the resulting CF1-CP, CF2-CP, and  
CF3-CP algorithms appear to solve their respective convex feasibility problems efficiently –  
an important feature for the ill-conditioned data model corresponding the limited angular-  
range scan. The efficiency is demonstrated by comparing convergence of CF1-CP with known  
algorithms for large-scale optimization. We then note that CF2-CP and CF3-CP, for which  
500 there is no alternative algorithm that we know of, appears to have similar convergence rates  
to CF1-CP.

Aside from the issue of efficiency, algorithm design can benefit from the different point  
of view offered by convex feasibility. For imaging applications this design approach extends  
naturally to considering non-convex feasibility sets [9, 24], which can have some advantage  
505 particularly for very sparse data problems. Future work will consider extension of the  
presented methods to the non-convex case and application of the present methods to actual  
data for CT acquired over a limited angular-range scan.

## ACKNOWLEDGMENTS

This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was supported in part by NIH R01 grants CA158446, CA120540, and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

## Appendix A: Pseudocode for CF3-CP

The pseudocode for CF3-CP appears in Fig. 9, and we explain variables not appearing in Secs. II A and II B. At Line 6 the symbol  $\nabla$  represents a numerical gradient computation, and it is a matrix which applies to an image vector and yields a spatial-vector image, where the vector at each pixel/voxel is either 2 or 3 dimensional depending on whether the image reconstruction is being performed in 2 or 3 dimensions. Similarly, the variables  $\mathbf{t}$  and  $\mathbf{z}_n$  are spatial-vector images. At Line 7 the operation “ $|\cdot|$ ” computes the magnitude at each pixel of a spatial-vector image, accepting a spatial-vector image and yielding an image vector. This operation is used, for example, to compute a gradient-magnitude image from an image gradient. The ratio appearing inside the square brackets of Line 7 is to be understood as a pixel-wise division yielding an image vector. It is possible that at some pixels the numerator and denominator are both zero in which case we define  $0/0 = 1$ . The quantity in the square brackets evaluates to an image vector, which then multiplies a spatial-vector image; this operation is carried out, again, in pixel-wise fashion where the spatial-vector at each pixel of  $\mathbf{t}$  is scaled by the corresponding pixel-value. At Line 8,  $\nabla^T$  is the transpose of the matrix  $\nabla$ , see Ref. [6] for one possible implementation of  $\nabla$  and  $\nabla^T$  for two dimensions.

The pseudocode for the function  $proj_{Diamond(\gamma)}(\mathbf{x})$  appears in Fig. 10. This function is essentially the same as what is listed in Figure 1 of Ref. [19]; we include it here for completeness. The “if” statement at Line 2, checks if the input vector  $\mathbf{x}$  is already in  $Diamond(\gamma)$ . Also, because the function  $proj_{Diamond(\gamma)}(\mathbf{x})$  is used with a non-negative vector argument in Line 7 of Fig. 9, the multiplication by  $sign(\mathbf{x})$  at the end of the algorithm in Fig. 10 is unnecessary for the present application. But we include this  $sign$  factor so that the function applies to any  $N$ -dimensional vector.

---

```

1:  $L \leftarrow \|X\|_2$ ;  $\tau \leftarrow 1$ ;  $\sigma \leftarrow 1/L^2$ ;  $n \leftarrow 0$ 
2: initialize  $\mathbf{f}_0$ ,  $\mathbf{y}_0$ , and  $\mathbf{z}_0$  to zero vectors
3:  $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$ 
4: repeat
5:    $\mathbf{y}_{n+1} \leftarrow \max(\|\mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})\|_2 - \sigma\epsilon, 0) (\mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g}))$ 
6:    $\mathbf{t} \leftarrow \mathbf{z}_n + \sigma\nabla\mathbf{f}_n$ 
7:    $\mathbf{z}_{n+1} \leftarrow \mathbf{t} \left[ (|\mathbf{t}| - \sigma \text{proj}_{\text{Diamond}(\gamma)}(|\mathbf{t}|/\sigma)) / |\mathbf{t}| \right]$ 
8:    $\mathbf{f}_{n+1} \leftarrow [\mathbf{f}_n - \tau(X^T\mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}} + \nabla^T\mathbf{z}_{n+1})] / (1 + \tau)$ 
9:    $\theta \leftarrow 1/\sqrt{1 + 2\tau}$ ;  $\tau \leftarrow \tau\theta$ ;  $\sigma \leftarrow \sigma/\theta$ 
10:   $\bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$ 
11:   $n \leftarrow n + 1$ 
12: until  $n \geq N$ 

```

---

**FIG. 9:** Pseudocode for  $N$ -steps of the accelerated CP algorithm instance for solving Eq. (10) with parameters  $\epsilon$  and  $\gamma$ . Variables are explained in the text, and pseudocode for the function  $\text{proj}_{\text{Diamond}(\gamma)}(\mathbf{x})$  is given in Fig. 10.

- 
- [1] R. Gordon, R. Bender, and G. T. Herman, “Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography,” *J. Theor. Biol.* **29**, 471–481, (1970).
  - 540 [2] G. T. Herman, *Image Reconstruction from Projections*, New York: Academic, 1980.
  - [3] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, “Implementation of an optimal first-order method for strongly convex total variation regularization,” *BIT Num. Math.* **52**, 329–356, (2012).
  - [4] M. Defrise, C. Vanhove, and X. Liu, “An algorithm for total variation regularization in high-  
545 dimensional linear problems,” *Inv. Prob.* **27**, 065002, (2011).
  - [5] S. Ramani and J. Fessler, “A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction,” *IEEE Trans. Med. Imag.* **31**, 677–688, (2012).



---

```

1: function  $proj_{Diamond(\gamma)}(\mathbf{x})$ 
2:   if  $\|\mathbf{x}\|_1 \leq \gamma$  then
3:     return  $\mathbf{x}$ 
4:   end if
5:    $\mathbf{m} = |\mathbf{x}|$ 
6:   Sort  $\mathbf{m}$  in descending order:  $m_1 \geq m_2 \geq \dots m_N$ 
7:    $\rho \leftarrow$  largest  $j$  such that  $m_j - \frac{1}{j} \left( \sum_{k=1}^j m_k - \gamma \right) > 0$ , where  $j \in [1, N]$ 
8:    $\theta \leftarrow (1/\rho) \left( \sum_{k=1}^{\rho} m_k - \gamma \right)$ 
9:    $\mathbf{w} = \max(|\mathbf{x}| - \theta, 0)$ 
10:  return  $\mathbf{w} \, sign(\mathbf{x})$ 
11: end function

```

---

**FIG. 10:** Pseudocode for the function  $proj_{Diamond(\gamma)}(\mathbf{x})$ , which projects  $\mathbf{x}$  onto the  $\ell_1$ -ball of scale  $\gamma$ . This function appears at line 7 of algorithm in Fig. 9. The vector  $\mathbf{x}$  is taken to be one-dimensional with length  $N$ , and the individual components are labeled  $x_i$  with index  $i$  being an integer in the interval  $[1, N]$ .

[6] E. Y. Sidky, J. H. Jørgensen, and X. Pan, “Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm,” *Phys. Med. Biol.* **57**, 3065–3091, (2012).

[7] P. Combettes, “The convex feasibility problem in image recovery,” *Adv. Imag. Electron Phys.* **95**, 155–270, (1996).

[8] P. L. Combettes, “The foundations of set theoretic estimation,” *IEEE Proc.* **81**, 182–208, (1993).

[9] X. Han, J. Bian, E. L. Ritman, E. Y. Sidky, and X. Pan, “Optimization-based reconstruction of sparse images from few-view projections,” *Phys. Med. Biol.* **57**, 5245–5274, (2012).

[10] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imag. Vis.* **40**, 120–145, (2011).

[11] J. H. Jørgensen, E. Y. Sidky, and X. Pan, “Quantification of admissible undersampling for sparsity-exploiting iterative image reconstruction in X-ray CT, ,” (2011), arxiv preprint

arxiv:1109.0629 (<http://arxiv.org/abs/1109.0629>).

[12] J. Nocedal and S. Wright, *Numerical Optimization, 2nd ed.*, Springer, 2006.

[13] C. C. Paige and M. A. Saunders, “LSQR: an algorithm for sparse linear equations and sparse least squares,” *ACM Trans. Math. Soft.* **8**, 43–71, (1982).

565 [14] R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, 1970.

[15] T. Pock and A. Chambolle, “Diagonal preconditioning for first order primal-dual algorithms in convex optimization,” in *International Conference on Computer Vision (ICCV 2011)*, 2011, 1762–1769.

[16] Y. Nesterov, “A method of solving a convex programming problem with convergence rate  $O(1/k^2)$ ,” in *Soviet Mathematics Doklady*, 27, no. 2, 1983, 372–376.

[17] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems,” *IEEE Trans. Imag. Proc.* **18**, 2419–2434, (2009).

[18] C. Vogel, *Computational methods for inverse problems*, Society for Industrial Mathematics, 2002, 23.

575 [19] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, “Efficient projections onto the  $\ell_1$ -ball for learning in high dimensions,” in *Proceedings of the 25th international conference on Machine learning*, ACM, 2008, 272–279.

[20] J. H. Jørgensen, P. C. Hansen, E. Y. Sidky, I. S. Reiser, and X. Pan, “Toward optimal X-ray flux utilization in breast CT,” in *Proceedings of the 11th International Meeting on Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2011, arxiv preprint arxiv:1104.1588 (<http://arxiv.org/abs/1104.1588>).

580 [21] I. Reiser and R. M. Nishikawa, “Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise,” *Med. Phys.* **37**, 1591–1600, (2010).

[22] J. H. Jørgensen, E. Y. Sidky, and X. Pan, “Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT,” in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, Valencia, Spain, 2011, 2640–2643.

[23] E. Y. Sidky and X. Pan, “Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization,” *Phys. Med. Biol.* **53**, 4777–4807, (2008).

590 [24] D. R. Luke, “Relaxed averaged alternating reflections for diffraction imaging,” *Inv. Prob.* **21**, 37–50, (2005).